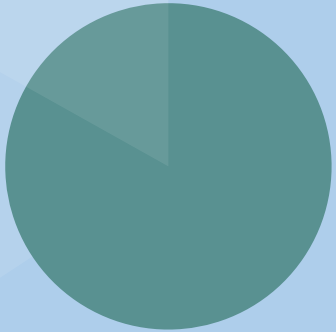
Apple apps vs Google apps

Did Apple Store apps receive better reviews than Google Play apps?



Apple vs Google: Process Overview

Steps taken to complete this project included:

1. Sourcing data from Kaggle and loading data into Python using pandas
2. Cleaning and transforming the original data into a relevant dataset by fixing the data, changing column names, eliminating irrelevant rows, and combining datasets by employing numpy
3. Visualizing data using matplotlib (ie boxplot, histogram) and analytically using scipy and random (ie normaltest, permutation)
4. Modeling by establishing null (observed difference due to chance) and alternative hypotheses (observed difference is due to platform type) and defining the significance level
5. Drawing a conclusion based on data analysis

Apple vs Google: Statistical Findings

Observed Difference

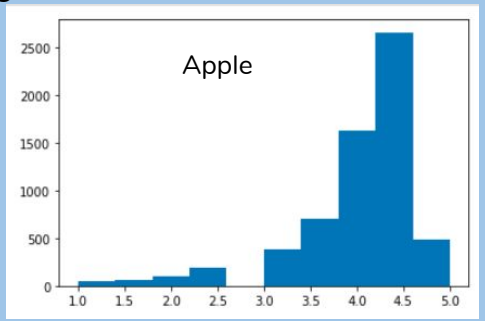
- Appears to be insignificant for ratings reviews of Google and Apple reviews
- Mean of Google apps rating - mean of Apple apps rating = 4.191757 - 4.049697 = 0.14206

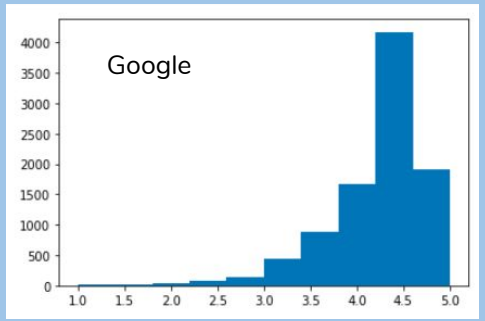
Hypotheses

- Null (H0): Platform rating is due to chance
- Alternate (HA): Platform rating is NOT due to chance
- Alpha/significance level: 0.05, a set threshold to interpret against p-value

Statistical Tests

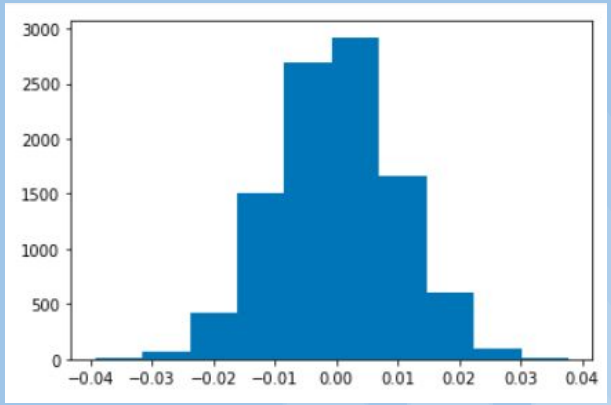
- Ratings are likely not normally distributed since p-values for Apple and Google are 0.0
- Confirms data is not normally distributed by looking at histograms for Apple and Google reviews





Apple vs Google: Testing hypothesis

- Permutation test, a non-parametric test, was used to test the null hypothesis
- Random permutation test shows mean difference of 0.001103--big difference to observed difference of 0.14206. Hence, our null hypothesis is false, platform does affect ratings?!
- To validate this is true, ran 10,000 permutation (shuffling the ratings) tests by calculating the mean ratings for Google and Apple apps and the difference between these for each one, and then take the average of all of these differences. Results of these differences are plotted on the histogram



Apple vs Google: Conclusion

Permutation tests show there were 0 number of times that differences in our difference list is at least as extreme as our observed difference of 0.14206054755. And if the count is less than 5% of the defined significance level, we therefore:

REJECT NULL HYPOTHESIS and ACCEPT ALTERNATE

HYPOTHESIS that platform rating is not due to chance